

# Experimental design and statistical evaluation of limiting dilution assays

Citation for published version (APA):

Strijbosch, L. W. G. (1989). *Experimental design and statistical evaluation of limiting dilution assays*. [Doctoral Thesis, Maastricht University]. Rijksuniversiteit Limburg. <https://doi.org/10.26481/dis.19890421ls>

## Document status and date:

Published: 01/01/1989

## DOI:

[10.26481/dis.19890421ls](https://doi.org/10.26481/dis.19890421ls)

## Document Version:

Publisher's PDF, also known as Version of record

## Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

## General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.umlib.nl/taverne-license](http://www.umlib.nl/taverne-license)

## Take down policy

If you believe that this document breaches copyright please contact us at:

[repository@maastrichtuniversity.nl](mailto:repository@maastrichtuniversity.nl)

providing details and we will investigate your claim.

## SUMMARY

This thesis is a contribution to the theory of Limiting and Serial Dilution Assays (LDA's, SDA's) which are used in many areas of bioresearch. In general, the aim of such an assay is the estimation of the frequency of a particular cell type among other but similar cells (LDA), or the estimation of the density of particular cells or microorganisms per unit volume of solution (SDA). In both LDA and SDA the unknown parameter is commonly estimated by using the "single-hit Poisson model" with binary (positive or negative) data produced by samples taken from different dilutions. In this study research has been performed on two important statistical issues of bioassays. These are the construction of an experimental design and the choice of the estimation techniques to be used.

For the setup of a design for a LDA or SDA one has to choose the number of dilutions, the number of replicates in each dilution and the (mean) number of cells/organisms per replicate. A design which has been constructed properly should take advantage of existing a priori information of the researcher and incorporate simple design parameters chosen by the researcher in such a way that the production of useless data is limited and the precision of the estimator is controlled. In Chapter 2 a useful and simple method for setting up a design on the basis of the researcher's criteria and a priori information about the unknown parameter is presented. Refinements for this design method and a program to perform these techniques are presented in Chapter 3. The methods of Chapter 2 and 3 regard the experimental design for LDA's and SDA's. The basic design method of Chapter 2 can be generalized to some other related and widely used biological models, such as the "logistic regression model" which is most common in bioassay. As the advance information on *two* unknown parameters have to be incorporated here, the experimental design is more complicated. The corresponding generalizations for three different cases of advance information are presented in Chapter 4.

The second main issue contained in this dissertation is the study of the commonly used estimation methods and some bias-reducing methods when applied to dilution analysis. In Chapter 5 two conventional statistical procedures (the Maximum Likelihood (ML) method and the Minimum Chi-square

(MC) method) are described. For an experiment with all positive responses the corresponding estimates are infinite. Thus, modified ML and MC estimators are proposed in order to reduce the bias of these estimators from infinity to the desired order. However, the resulting estimators are still biased. Little is known about bias and efficiency of the modified ML and MC estimators when applied to data from dilution experiments with small or moderate size. Resampling methods like the jackknife and the bootstrap provide estimators of bias and variance for an extremely wide class of statistics. In general, these methods are applied to independent identically distributed (iid) random variables. There are also several reasons to consider some jackknife and bootstrap versions based on non-iid variables. In Chapter 6 three jackknife and four bootstrap versions of the ML method are discussed and applied to dilution analysis.

Results on Monte Carlo comparison of estimators have been published by several authors. However, there is a difficulty in comparing these results because of the absence of generally accepted design methods. Some controversies emerged partly as a consequence of this kind of difficulty. Research on estimators in dilution analysis requires a design method that can be used as a frame of reference. Thus, the design method proposed in Chapter 2 is used for the simulation study conducted in this dissertation. A PASCAL-program has been written to perform the Monte Carlo experiments for the comparison of the statistical methods described in Chapters 5 and 6. The simulation results are presented and discussed in Chapter 7. The criteria used to compare the estimation methods concern the bias, the coefficient of variation (based upon the mean square error) and the realized confidence level (i.e. the proportion of times that the realized 95% confidence interval contains the true value of the parameter). The main results can be summarized as follows. The corresponding jackknife and bootstrap versions of the ML estimator are very strongly related. Generally, all resampling methods reduce the bias of the ML estimator significantly to values around zero. Provided that the design used is not too small, the most preferable method is a particular jackknife version (obtained by deleting one individual replicate culture at a time). The MC estimator is shown to be inferior to the ML estimator.

# SAMENVATTING

## EXPERIMENTEEL ONTWERP EN STATISTISCHE ANALYSE VAN BIOLOGISCHE VERDUNNINGSREEKSEN

In dit proefschrift wordt een bijdrage geleverd aan de theorie over biologische verdunningsreeksen. In de literatuur worden twee typen verdunningsreeksen onderscheiden, te weten de "Limiting Dilution Assay" (LDA) en de "Serial Dilution Assay" (SDA). Bij LDA gaat het er veelal om te schatten welke proportie van een verzameling gelijksoortige cellen een bepaalde eigenschap bezit, terwijl bij SDA de dichtheid van een bepaald type cel per volume eenheid wordt geschat. Zowel bij een LDA als een SDA wordt bij het schatten van de onbekende parameter meestal gebruik gemaakt van de veronderstelling dat het aantal relevante cellen per volume eenheid een Poisson verdeling heeft, en dat de aanwezigheid van één relevante cel in een volume eenheid al met zekerheid kan worden vastgesteld. Aangezien de te schatten parameter onbekend is en het tellen van cellen doorgaans onmogelijk is, dient met verschillende verdunningen gewerkt te worden ten einde onafhankelijk van de waarde van de parameter in een bepaald gebied een schatting te kunnen bepalen. In het hier beschreven onderzoek worden zowel de experimentele proefopzet als de analyse van verdunningsreeksen bestudeerd.

Voor een LDA of een SDA dient bij het ontwerp van een experiment het aantal verdunningen gekozen te worden, alsmede het aantal te herhalen bepalingen per verdunning en het (gemiddeld) aantal cellen per herhaling. Een goed ontwerp houdt rekening met de voorkennis van de onderzoeker en bevat ontwerp parameters die de statistische eigenschappen van het ontwerp, zoals de te bereiken nauwkeurigheid, bepalen. In hoofdstuk 2 wordt een dergelijk ontwerp gepresenteerd, terwijl in hoofdstuk 3 diverse verfijningen worden uitgewerkt en een computer-programma wordt beschreven dat deze technieken uitvoert. De basisprincipes uit hoofdstuk 2 kunnen worden gegeneraliseerd voor toepassing op andere biologische modellen (met twee onbekende parameters) die sterk aan LDA's en SDA's gerelateerd zijn, zoals het "logistische regressie model". Aangezien hier de voorkennis met betrekking tot *twee* onbekende parameters in het experimenteel ontwerp verwerkt moet worden, is een dergelijk ontwerp ingewikkelder. In hoofdstuk 4

worden voor drie verschillende gevallen van aanwezige voorkennis ontwerp voorstellen gedaan.

Het tweede hoofd onderwerp in dit proefschrift betreft de bestudering van de gebruikelijke schattingsmethoden en enige onzuiverheid-reducerende schattingsmethoden voor toepassing bij verdunningsreeksen. In hoofdstuk 5 worden twee conventionele schatters besproken, namelijk de "Maximum Likelihood" (ML) schatter en de "Minimum Chi-square" (MC) schatter. Bij een experiment met alleen positieve uitslagen zijn de corresponderende schattingen oneindig. Ten einde de onzuiverheid van deze schatters terug te brengen tot de gewenste orde, worden gewijzigde ML en MC schatters voorgesteld. Deze schatters vertonen echter nog steeds een onzuiverheid, die bij experimenten van kleine of bescheiden omvang bezwaarlijk kan zijn. Vanwege de brede toepasbaarheid van zgn. jackknife en bootstrap methoden kan worden geprobeerd de onzuiverheid te schatten om daarmee een schatter te construeren die minder onzuiver is. Jackknife en bootstrap methoden verschaffen ook automatisch een schatter voor de variantie. In het algemeen worden deze methoden toegepast op onafhankelijke en identiek verdeelde stochastische variabelen. De speciale toepassing op verdunningsreeksen echter suggereert ook varianten die zijn gebaseerd op onafhankelijke maar niet identiek verdeelde variabelen. In hoofdstuk 6 worden drie jackknife en vier bootstrap varianten van de ML methode besproken en toegepast op verdunningsreeksen.

Het ontbreken van algemeen geaccepteerde ontwerpmethoden voor verdunningsreeksen maakt het moeilijk om de Monte Carlo resultaten van verschillende onderzoekers te vergelijken. Sommige tegenstrijdigheden in de literatuur kunnen hieraan worden toegeschreven. Onderzoek naar de eigenschappen van schatters voor de onbekende parameter in verdunningsreeksen vraagt om een ontwerpmethode, die kan worden gebruikt als een referentiekader. De in hoofdstuk 2 beschreven methode is hiervoor geschikt en wordt gebruikt in het simulatie onderzoek, dat in hoofdstuk 7 beschreven is. Dit onderzoek vergelijkt m.b.v. een speciaal voor dit doel in PASCAL geschreven programma de volgende drie eigenschappen van de in hoofdstukken 5 en 6 beschreven schatters: de gemiddelde relatieve onzuiverheid, de variatiecoëfficiënt (afgeleid van de gemiddelde kwadratische fout) en de proportie van het aantal keren dat het gerealiseerde 95% betrouwbaarheidsinterval de werkelijke waarde van de

parameter bevat. De belangrijkste resultaten kunnen als volgt worden samengevat. De met elkaar corresponderende jackknife en bootstrap varianten hebben sterk op elkaar gelijkende eigenschappen. In het algemeen reduceren ze de onzuiverheid van de ML schatter significant tot waarden rond nul. Mits het ontwerp een voldoende aantal elementaire bepalingen bevat kan worden gesteld dat de jackknife variant die wordt verkregen door het beurteilings weglaten van één elementaire bepaling de voorkeur geniet. Van de MC schatter wordt aangetoond, dat deze niet te verkiezen is boven de ML schatter.